



**Manchester
Metropolitan
University**

Vega, J, Couth, S, Poliakoff, E, Kotz, S, Sullivan, M, Jay, C, Vigo, M and Harper, S (2018) Back to Analogue: Self-Reporting for Parkinson's Disease. In: CHI Conference on Human Factors in Computing Systems (CHI '18), 21 April 2018 - 26 April 2018, Montreal, Canada.

Downloaded from: <https://e-space.mmu.ac.uk/620850/>

Publisher: ACM

DOI: <https://doi.org/10.1145/3173574.3173648>

Please cite the published version

<https://e-space.mmu.ac.uk>

Back to Analogue: Self-Reporting for Parkinson's Disease

Julio Vega*, Samuel Couth†, Ellen Poliakoff‡, Sonja Kotz‡, Matthew Sullivan§, Caroline Jay*, Markel Vigo*, Simon Harper*

School of Computer Science*
School of Biological Sciences†
University of Manchester, UK
first.lastname@manchester.ac.uk

Department of
Neuropsychology‡
University of Maastricht, NL
sonja.kotz@maastrichtuniversity.nl

School of Science and the
Environment§
Manchester Metropolitan
University, UK
m.sullivan@mmu.ac.uk

ABSTRACT

We report the process used to create artefacts for self-reporting Parkinson's Disease symptoms. Our premise was that a technology-based approach would provide participants with an effective, flexible, and resilient technique. After testing four prototypes using Bluetooth, NFC, and a microcontroller we accomplished almost full compliance and high acceptance using a paper diary to track day-to-day fluctuations over 49 days. This diary is tailored to each patient's condition, does not require any handwriting, allows for implicit reminders, provides recording flexibility, and its answers can be encoded automatically. We share five design implications for future Parkinson's self-reporting artefacts: reduce participant completion demand, design to offset the effect of tremor on input, enable implicit reminders, design for positive and negative consequences of increased awareness of symptoms, and consider the effects of handwritten notes in compliance, encoding burden, and data quality.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords

Self-report; Parkinson's Disease; Paper diary; EMA

INTRODUCTION

Parkinson's Disease is a neurodegenerative disease with a wide variety of motor and non-motor symptoms that fluctuate within hours or days. People with Parkinson's (PwP) are commonly assessed every six months, which makes it difficult to monitor disease progression and thus tailor medication to the fluctuations of their condition. As part of a parent project designed to monitor Parkinson's 24/7 for one year, we have conducted a long-term study of different artefacts for collecting daily feedback about symptoms, which will be used as ground truth in the parent investigation. Such a child study,

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI 2018 April 21–26, 2018, Montreal, QC, Canada

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5620-6/18/04.

DOI: <https://doi.org/10.1145/3173574.3173648>



Figure 1. First prototype using Flic buttons on an A4 piece of paper

which we report here, culminated in an evaluation of what emerged as the optimal device: a paper diary.

First, we provide an overview of our parent project for context. Every participant uses a smartphone loaded with the Aware Framework app that collects inertial, location, environmental, and interaction data 24/7 [11]. We aim to find a set of metrics inferred from smartphone data that is correlated with the clinical progression of the disease, i.e. a digital biomarker.

Using these digital biomarkers, we expect to measure longitudinal changes in people's symptoms as well as within-day and day-to-day changes. Traditional clinical scales like the Unified Parkinson's Disease Rating Scale (UPDRS) [15] will be used to verify the long-term fluctuations, but self-reporting is our best option to validate daily changes. We could have carried out interviews or electronic/analogue assessments like spiral drawing [43] but they are time-consuming and disruptive if we consider our required recording frequency and the duration of the study. What is more, these individual tests measure a narrow set of symptoms, and thus we would need to use a battery of tests to cover people's conditions which would increase patient burden. Consequently, we rely on people's ability to report the severity and impact of their symptoms as they do when completing part II of the UPDRS or questionnaires like the Parkinson's Disease Questionnaire (PDQ-39) [42] every six months or longer.

However, self-reporting using an electronic or a paper diary presents its own set of challenges that affect user compliance, acceptance, and data quality. Indeed, a digital approach can be more suitable for precise timing and within-day assessments [16, 27, 50], whereas an analogue approach is convenient for once-a-day assessments [50, 27]. Since we can monitor our

participants at both time scales, our goal was not to compare a digital and an analogue method but rather design different prototypes, starting with an electronic button-based diary and going back to analogue when we found out a paper diary was an accessible, frictionless, personalised, portable, low-burden, and automatic encoded approach. Thus, our contributions in this paper are the following:

1. An open source pen and paper diary that reached almost full compliance and high acceptance (see Paper Diary).
2. Five design implications for Parkinson's self-reporting artefacts: reduce participant completion demand, design to offset the effect of tremor on input, enable implicit reminders, design for positive and negative consequences of increased awareness, and consider the effects of handwritten notes in compliance, encoding burden, and data quality (See Discussion).

BACKGROUND

Ecological Momentary Assessment (EMA) is a collection of methodologies to sample participants' behaviour and experiences in naturalistic settings multiple times over time [48]. EMA is not delimited to a particular technology or methodology but includes the use of paper and electronic diaries which have been used extensively in clinical, psychological, and computer science research. Traditionally, analogue diaries use pen and paper to record answers to open or closed questions while digital diaries can be implemented in different hardware: smartphones (or their predecessors Personal Assistant Devices), desktop computers, laptops, and wearables.

Although ours is not a traditional diary study (its goal is to collect ground truth), we frame our contributions using the literature of diary design as the goal of both approaches is the same: to capture participants' feedback over time. Both electronic and paper diaries have advantages and disadvantages. Paper diaries are cheap, portable, robust, can be pictured, memorised and scanned with ease [30], require less start-up time [12], and are easy for people to use [4]. On the other hand, for some diary designs, it can be difficult to evaluate the validity of people's answers (more relevant to some research areas than others), participants can have privacy concerns as diaries can be read by anyone, and there is a risk of forgetfulness or retrospective error when filling in missed entries with truthful data or fabricated answers. Moreover, these diaries can also be resource consuming and error-prone when it comes to data entry, encoding, and analysis [4].

In contrast, electronic diaries can alleviate some of these issues and offer extra features. For example, they can have richer user interfaces with wider response formats and presentation styles [4, 37], provide a standardised test environment, reach out to larger user populations, enable online scoring, and improve the quantity and quality of writing [37]. Researchers can also evaluate the compliance of the study protocol thanks to automatic timestamps and signalling [4, 49], prevent retrospective completion [49], reduce data management burden (entry, encoding, transfer, storage, analysis) [4, 25, 27, 37], and adapt the diaries dynamically to the provided responses [4]. Nevertheless, as with any other computer-based task,

their hardware or software can malfunction, and the cost of replacing or fixing it can be high, the characteristics of their screen (if any) can pose difficulties to certain populations (i.e. people with motor or visual impairment), and there is still a concern for data confidentiality especially if data is transmitted over the Internet [37]. Additionally, there is a higher financial commitment than paper diaries (although this might offset the cost of manual data management using paper), training might be required depending on the device and participants, there could be an overhead for researchers as setting them up may require high technical skills [4, 49], and electronic diaries implemented in devices like smartphones rely on batteries that can limit certain study designs. What is more, there is a risk of perpetuating a "digital divide", i.e., conducting research with more convenient or accessible populations where high-tech devices are easy to deploy or where computer literacy may affect participants' willingness to enrol [4].

From the previous paragraphs, it might seem evident that electronic diaries have a distinct advantage over paper diaries. Indeed, Dale et al. [10] in a systematic review of the use of Personal Device Assistants vs paper diaries in randomised and quasi-randomised controlled trials suggested that there might be an element of publication bias, because all the reports included, and almost all those excluded, were all in favour of electronic diaries. This might also be exacerbated by an unfair comparison of paper and electronic diaries [50] and by some research misinterpreting or generalising research focused on the attributes of a diary sub-type and sampling strategy [7], thus leading to strong recommendations like: "in our view, computerised methods are always preferable" [9]. In fact, comparisons between electronic and paper diaries have had mixed results suggesting that both types yield data of comparable quality, that there is an advantage to electronic diaries, or even that neither paper nor electronic diaries provide data of sufficient accuracy to serve as a measure of outcome in clinical trials [50].

Unsurprisingly, evidence for participants preference for either method is mixed too. People expressed approval for electronic diaries but did not think any was onerous [24], others found a slight majority of participants that preferred electronic devices [12] or a significant majority that favour digital approaches [18, 23, 25, 33, 53]. In any case, there is a consensus that more research is needed to analyse the comparability, compliance, data quality and user acceptance of electronic and paper diaries [5, 7, 27, 50].

So, how should researchers choose one method or the other? In the context of electronic and paper-based tasks, Noyes et al. recommended that choosing one or the other should be based on their advantages and disadvantages and their merits in relation to task demands and required performance outcomes [37]. In fact, it is suggested that the effectiveness of a diary study depends on careful consideration of the research questions to answer [4] and the appropriateness of either method to tackle them [16]. This is more evident after considering the many factors that can influence user preference and compliance: study design (i.e., having unreasonable expectations of patients [21]), participant motivation [16, 36], population

demographics [36], participants familiarity with technology [27, 50], factors inherent to the reported phenomena (like reluctance to report relapses in discharged alcohol-dependant participants [5]), and factors related to disease self-monitoring (like the fear of damaging the relationship with healthcare professionals, termination of access to medication, or discomfort monitoring indicators of pathology [7]).

In the context of Parkinson's, diaries have been used to monitor falls and motor/non-motor symptom fluctuations. Fall diaries are the accepted gold standard for fall quantification [14] and normally, they require patients to record the time of a fall beside its circumstances using open-ended questions after each episode [2, 20, 31]. Since these diaries are event-based, they can represent a lighter burden for PwP compared to fixed time diaries and therefore be suitable for long-term monitoring (e.g. six months [2], one year [31], or even four years [20]). Nevertheless, Hunter et al. observed attrition in diary usage (51% after 48 months) mainly related to study withdrawal and non-compliance [20].

Motor fluctuation diaries, on the other hand, are meant to measure changes between ON and OFF time (when medication is and is not working, respectively). Therefore, they require participants to log their motor status choosing among 3, 4, or 5 different states, every 30 min for 2 to 7 days [17, 44]. Although they are recommended [1] and accepted as endpoints for measuring the efficacy of Parkinson's medication, recall and diary fatigue can occur and compliance can be challenging [40]. Indeed, Lyons et al. suggest "it does not appear to be a realistic expectation that the majority of patients will complete each entry at half-hour intervals for 7 consecutive days" [32]. Following a similar approach, the SCOPA-DC diary was created to measure Parkinson's symptoms fluctuations seven times a day [34]. Although also demanding, this diary had good internal consistency and reproducibility in stable patients over two weeks, and there was evidence of its construct validity. Finally, Nyholm et al. compared the use of electronic (using Portable Device Assistants) and paper diaries for motor functioning asking people to complete 10-11 questions every two hours on two nonconsecutive days per week over four weeks [39]. Although they did not expect high compliance with the paper diary, they found that 78% participants completed the diary within a 15-minute window after the scheduled time and 98% did so when no limit was considered; they attributed these numbers to user's motivation or invalidly completed diaries (prospectively or retrospectively).

The identified Parkinson's diaries have focused on patients with motor fluctuations and dyskinesias (involuntary movement) using event- or time-based strategies. What is more, for the latter "the extent to which accurate diary data can be collected beyond a time period remains unknown" [40]. Consequently, since the design of a diary must be tailored to the study and the use of such method has not been reported for longitudinal, end-of-day, day-to-day assessments of symptoms, the insights and design implications that we share represent a valuable contribution. In the rest of this paper, we detail the agile iterative design process that led us to create a paper diary guided by our participants feedback, constrained by our

study's needs, and taking into account the four suggestions from Bolger et al. [4] to implement analogue diary studies: 1) make the diaries easily portable, 2) reduce the possibility of participant error, 3) pilot test the diaries on participants from the population to be studied, and 4) maintain ongoing contact with participants, in a personal yet unobtrusive manner.

FROM A DIGITAL TO AN ANALOGUE SYMPTOM DIARY

Over the course of seven months, a group of PwP and the research team co-designed six different prototypes to self-report Parkinson's symptoms. These included the use of Bluetooth, NFC and a microcontroller to collect within-day changes to finally produce an analogue diary that is more suitable for day-to-day fluctuations, does not interfere with the use of our monitoring tool (smartphone), and has significant advantages for our parent monitoring study.

We recruited seven (3 female) non-demented (ACE-III[19] > 88) and non-depressed (GDS[55] < 6) participants with mild to moderate Parkinson's Disease (Hoen & Yahr 1 and 2), age 66 ± 6.4 , years since diagnosis 5.8 ± 3.6 , daily dose of levodopa 335.7 ± 165.1 . We ask the reader to refer to the timeline in Figure 2 throughout this paper to have a better understanding of the development of this project.

We recruited P1 (Point A in Figure 2) to develop our first prototype. At first, we aimed to record within-day symptom fluctuations because smartphone data collected continuously 24/7 would allow us to track participant's daily activities at an hourly rate. Likewise, since we are using people's routines as a proxy to their condition, we asked participants for the overall impact of their disease on their habits rather than rating the severity of individual symptoms. Symptoms' impact is used as an end-point in research (i.e. PDQ-39) and is also important in Parkinson's management as medication tailoring is influenced by the participants' perception of their symptoms. Therefore, we agreed on the following surveying question "Overall, how have your symptoms impacted your day so far?" with answers based on a 4-point scale: "No impact", "Low impact", "Moderate impact" and "High impact". We chose an even number of steps because people tend to go to the middle of the scale if it has an odd number of points [51], and we had four points because P1 considered it would be difficult for participants to discern 6 or 8 levels of impact. Although this assessment question has not been clinically validated, it is suitable for obtaining an individual longitudinal overview as we will track personal fluctuation rates instead of performing a cross-sectional analysis.

We then looked for an electronic device to implement this prototype as electronics are suitable for precise timing and within-day assessments [7, 16, 27]. Such a device should have the following attributes to fit our study's requirements:

- Independent. It needed to be a separate entity from our monitoring tool (smartphone); using a mobile app would remind people that they were being monitored and therefore be prone to change their behaviour [35].
- Accessible. It needed to be usable by people with dexterity problems caused by tremor, bradykinesia (slowness of

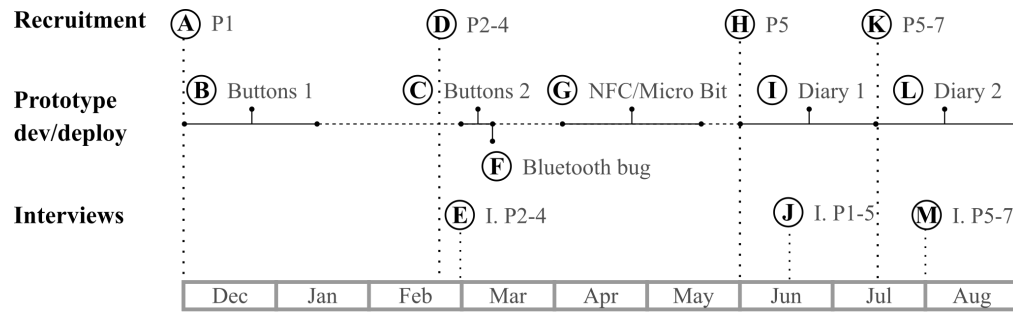


Figure 2. Timeline of relevant events in our design process. They include participants recruitment, prototypes development and deployment, and interviews with participants from December 2016 to August 2017.

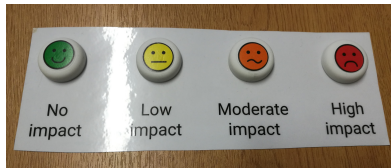


Figure 3. P1 trimmed down the A4 sheet of paper into a 20 x 7 cm piece where the buttons were held to make them portable

movement), dyskinesias, rigidity and other motor symptoms of Parkinson's.

- Frictionless. Since the main monitoring study is 12 months long, participants did not have to wait, configure or perform actions that might make self-reporting feel like a chore when they wanted to use the device.

First button-based prototype

For this prototype we used a set of 4 Bluetooth buttons from Flic¹ (Point B in Figure 2). They cost \$25 USD per unit and allow the user to register clicks, double clicks, and holding presses on the Flic mobile app using Bluetooth Low Energy (BLE). Any of these three actions can trigger a catalogue of predefined functions in the phone. Also, whenever a click happens a sound (beep) notifies the user, and when there is a problem with the connection, a red light glows in the face of the buttons. These buttons complied with the three design attributes mentioned above. Physically, they were a self-contained device, P1 thought people with moderate motor symptoms (Hoehn and Yahr scale ≤ 2) could use the buttons and, after pairing them, they would reconnect automatically allowing users to press them at any time. Each button represented one step in our severity scale (Low to High) and was glued to an A4 piece of paper where the surveying question was printed. We added four stickers to each button imitating a visual analogue pain scale and configured the Flic app to record each click in a Google Docs spreadsheet (Figure 1).

P1 tested this prototype during six weeks along with a Moto X Play smartphone running Android 5.1.1. Our only recommendations were he could keep the buttons anywhere, and he had to report his symptoms at least once a day. After the testing period was over, P1 suggested three critical things to address:

¹<https://flic.io/>

Portability

P1 emphasised he wanted a portable device. He reported the buttons were “easy to use” and that this made him keen to press them multiple times a day, which in turn meant he wanted to bring the board outside his house. Nonetheless, he could not do it because the rig was too big and the buttons could be pressed by accident in his bag. As a result, he trimmed down the A4 sheet around the buttons area (Figure 3):

“It’s half my day [being at work], I thought it was important to record data during the day as well.” (P1)

Personalised Monitoring

P1 suggested using four buttons to track four specific symptoms using 1 to 4 clicks to report their severity instead of asking participants for an overall view. He reasoned that in practice he only had a small set of symptoms that impacted his daily routine with different severity and at various times.

“When you are feeling bad you want to say what symptom is particularly affecting you now.” (P1)

Focus on daily experience

P1’s final remark was about rephrasing the surveying question. He thought it could be leading participants to report their symptoms had no impact as some people would have already adapted to their condition and what is affected now is not their daily routine per se but how people experience it.

“What’s happened is you’ve already allowed your symptoms to impact your day just by having a different lifestyle, so your symptoms used to impact your day [but not anymore] maybe experience is a better word.” (P1)

Second button-based prototype

For our next prototype, we made three changes to the button set following P1’s feedback (Point C in Figure 2). First, we placed the buttons inside a spectacles case to address the portable aspect. Second, we asked participants for their top 4 symptoms and assigned each to a single button to personalise the monitoring. Third, we rephrased the surveying question to: “How are your symptoms impacting your daily experience so far?” (Figure 4) to focus on daily experience which is used in clinical scales like the PDQ-39.

We recruited P2, P3, and P4 (Point D in Figure 2) and interviewed them to validate these changes (Point E in Figure 2).

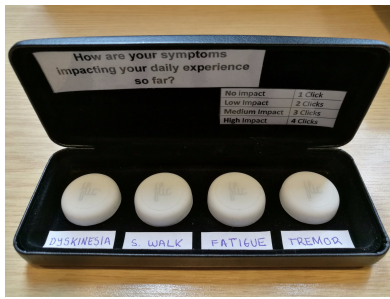


Figure 4. Second prototype using Flic buttons and one button per symptom. Spectacles case is 155 x 60 x 27 mm

We showed them one button without additions or modifications, explained its functionality, and then asked: “If you had one or more of these buttons, how would you use them to report the severity of your symptoms?”. From their responses all three backed up the idea of personalised monitoring:

“If those were the designated buttons, I would have one for each of my different symptoms. That would be much more sophisticated than just having a single button because it would differentiate what I’m communicating. In terms of severity, you want to keep things simple so you know one press for a bit, two presses for moderate and three presses for this is really awful.” (P3)

Finally, after showing them the finished second prototype, they praised its portability and simplicity:

“I’m completely interested in how you formulated this study I think it is very splendid. I mean, the golden bullet is simplicity, isn’t? The simpler you keep things the more effective. That’s thinking. This is a glasses case, isn’t it? Perfect! That’s really good, I love it.” (P3)

We deployed the second prototype to P1-4 for six weeks. However, due to a Bluetooth sensing bug (see Appendix, Project Constraints), we had to upgrade all participants’ phones to Android 6.1. In turn, the update caused the buttons to start working erratically after one week. (Point F in Figure 2). This was due to a different software problem² in Android 6.0 or higher that occurs when multiple equal BLE devices connect to certain Bluetooth chips. A patch for this bug is already available³, but is not merged into Android’s source code, thus prompting us to design a substitute tool since buying new phones was not possible. Nevertheless, even if this bug was fixed, we found out that our paper diary perfectly meets the needs and constraints of our monitoring study.

NFC cube

Our first alternative was a cube with 6 NFC tags attached to each side (Figure 5) developed at Point G of Figure 2. NFC tags are circuits that can store data and be read by a smartphone. We personalised this prototype to people’s symptoms in two ways: one where each tag represents a symptom and participants need to scan a tag multiple times in a row to report its severity, and the second, where three tags describe different



Figure 5. Alternative prototype using a wooden cube with a NFC tag on each face. The cube measures 6 cm on each side

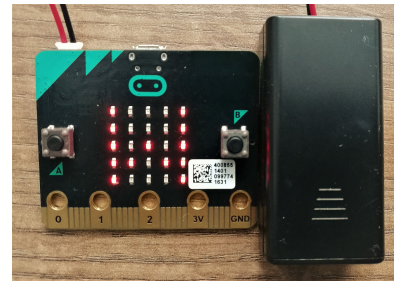


Figure 6. Alternative prototype using a Micro:bit. The black box on the left is the battery case that powers the device. Measurements: 5.5 x 7 x 1.5 cm.

symptoms and the other three represent three levels of severity. P1-4 agreed that the first option would be more intuitive because they would not need to fiddle with the cube, thus being more accessible. However, even though the cube was portable (6cm on each side), we discarded it because it was not independent or frictionless as, by Android’s security model, participants would need to unlock their phone to scan the tags and thus be reminded of the presence of the monitoring tool.

Micro:bit

Next, we built another prototype (Figure 6) using a Micro:bit at Point G of Figure 2. The Micro:bit is a microcontroller that has two buttons, a matrix of LEDs as a screen, and BLE capabilities, among other sensors and connections not relevant to this study. We made it personalised by programming the device so participants would press Button A (left) to report four symptoms labelled A to D, and button B (right) to select the intensity of that symptom (1 to 4). From a participant’s perspective, the device would be independent despite transmitting the selection to an Android application on the user’s smartphone. However, we did not go forward with this approach because we would have needed to 3D print a case to make it portable and we did not consider it frictionless as batteries would have to be swapped every four weeks and the device had issues reconnecting to the Moto X Play using BLE. Time was a limiting factor for us (see Appendix, Project Constraints), but it might be worth exploring this option further under different circumstances.

Paper Diary

Finally, we went back to analogue and decided to use pen and paper for three reasons: 1) paper diaries are suitable for

²<https://issuetracker.google.com/issues/37115854>

³<https://android-review.googlesource.com/#/c/315988/>

Tuesday, 30 May 2017 P01

How are your symptoms impacting your daily experience so far?

HH	MM	SS	AM/PM	Symptom	None	1	2	3	4	High
11	12	1	00	Dyskinesia						
10	am	2	15	Slow Walk						
9	pm	3	30	Fatigue						
8	pm	4	45							
7	6	5	45							

HH	MM	SS	AM/PM	Symptom	None	1	2	3	4	High
11	12	1	00	Dyskinesia						
10	am	2	15	Slow Walk						
9	pm	3	30	Fatigue						
8	pm	4	45							
7	6	5	45							

HH	MM	SS	AM/PM	Symptom	None	1	2	3	4	High
11	12	1	00	Dyskinesia						
10	am	2	15	Slow Walk						
9	pm	3	30	Fatigue						
8	pm	4	45							
7	6	5	45							

HH	MM	SS	AM/PM	Symptom	None	1	2	3	4	High
11	12	1	00	Dyskinesia						
10	am	2	15	Slow Walk						
9	pm	3	30	Fatigue						
8	pm	4	45							
7	6	5	45							

1 Please, fill out at least one row per day SKP

Figure 7. Diary’s first layout. Each page has four entries, and each entry has the top 3 symptoms for that participant which can be scored on a scale from 1 to 4 (None to High) depending on the impact of that symptom in people’s daily experience.

once-a-day assessments [50, 27] and our smartphone data (inertial, location, environmental, etc.) would still yield metrics relevant for this time scale, 2) our participants were familiar using diaries and notebooks which could ease adoption and 3) it would be a stand-alone, cheap, reliable and fast artefact to develop and deploy. Consequently, we created the layout showed in Figure 7 informed by our three design constraints for a self-reporting tool: independent, accessible and frictionless, and based on the feedback we got from the first two button prototypes, from which we knew our participants valued personalised monitoring and a portable design (Point I in Figure 2). Also, we anticipated that manual entry would be a burden for the research team and knew that, although for some types of paper diaries user compliance decreases over time [8], reducing participant demand might increase it [7, 16]. Our rationale to comply with these requirements and overcome possible drawbacks of analogue approaches was:

1. Independent. Not an issue, a paper diary is completely autonomous from our users’ smartphones.
2. Accessible. Participants do not have to write anything down. All input fields can be completed by colouring in the corresponding section.
3. Frictionless. Having a page per day and a bookmark (ribbon attached to the back cover) helped participants to find the section they needed to report on. We also assumed this would decrease the chance for input errors as suggested in [4]. On the recommendation of an external PwP with a hand tremor, we included a gel ink pen which could be clipped to the booklet’s spine, was retractable so that participants with dexterity issues did not have to deal with pen caps, and was anecdotally easier to write with.
4. Personalised. Participants choose their top 3 symptoms to be tracked. We reduced the number of symptoms from 4 to 3 and put them into four entries per page due to space

constraints and to P1’s opinion that it was not feasible to answer the diary more times than this.

5. Portable. We printed 60 pages on double-sided A5 booklets. We chose this size because some participants already had similar diaries and 60 pages because there were approximately 54 days between our assessment visits and thus we can replace a used diary with a new one at that point. This attribute is recommended too in [4].
6. Reduced demand. To lessen the burden of completing the diary, we went along with our original request to complete the diary only once a day, but we gave them three more entries to do it more times if they wished. We reinforced this notion by printing the bottom three rows with 50% less ink. This way we will have a log of day-to-day changes and possibly the most relevant within-day changes if our participants decided to do so.
7. Blending analogue and digital. We augmented our diary so it can be encoded automatically in an effort to tackle cumbersome data entry which is a drawback of analogue approaches. The diaries are scanned into TIF files and interpreted automatically by a custom Python web app using an Optical Mark Recognition algorithm similar to those employed in exams marking. The software to create, date and encode paper diaries is open source [52].

P1’s opinion on this diary iteration highlighted its attributes:

“This is much more likely to be completed. I think there are fewer barriers; I think it could be down at any time of the day, it gives me the option to provide more detail in terms of time points if I want to. I think it just hasn’t got the burden and the barriers of the technology, there’s too many things going on with the technology. The diary is less..., threatening is not the right word, that’s too strong, it’s less intimidating.” (P1)

We recruited P5 and deployed the diary to P1-5 (Point H in Figure 2). After two weeks we followed up with a phone interview inquiring about their general opinion on the diary’s design, surveying question, usage patterns, and what they liked and disliked about it (Point J in Figure 2). During the phone calls two aspects came to our attention: P2-4 reported that although they could perceive the severity of their symptoms changing during the day, its impact was harder to quantify, and either was something that they have adjusted in the past or something that they would only notice on the rare occasion when their symptoms were at their worst. To cope with it, they started to log the severity instead of the impact of their symptoms, so we adopted this format for the rest of the study.

Secondly, P5 decided to withdraw from the study because the diary felt like a chore and it was overwhelming for her to try to remember what she did during her day. She tried to get on with it but the prospect of completing the diary while on holiday was the tipping point.

“It ends up being filled in retrospectively because I don’t take it everywhere with me so then I have to remember where I was and what was happening earlier today. For instance, if I was walking somewhere then I need to be

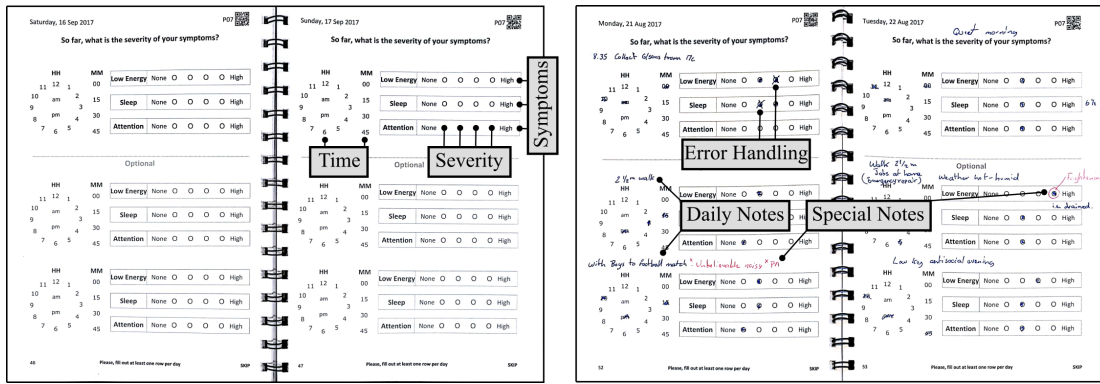


Figure 8. Left: Diary’s second layout printed as a double-sided A5 booklet with three sections to log time and the severity of 3 symptoms on a 4 point scale. Right: note scheme used by P7 with daily handwritten notes and extra notes for exceptional circumstances during his day; there is also an example of error handling by the participant, crossing out an undesired rating

specific about the time because you will need to match it with what your computer says was happening at that moment and quite often I don’t remember in that kind of detail so I’m kind of guessing.” (P5)

However, during our interview it was clear that we had a miscommunication problem, as she assumed we were matching the diary with her daily life and therefore, she had to record the time, symptoms’ severity and even notes on the margin about her activities as precisely as possible. Such was the case that four weeks later, she agreed to re-enroll after clarifying the goals of the study. The workload that she put in practice might be comparable to that of the diaries meant to capture medication cycles, where researchers looking for 30 minutes or 1 hour spaced logs do not ask participants to use the diary for more than 2-7 days within a week.

“I’m assuming that you’re trying to match what I say with what is on the tracking thing. So that I need to be as specific as I can about the time [...], I just assumed that you would be trying to put the two things together.” (P5)

In any case, after P1-4’s feedback and P5’s withdrawal, we made three minor changes to the diary for our next 6-week deployment (Figure 8). First, we included three entries per page since no participant filled in more than that on any day. Second, we marked the last two entries of each page as “Optional” complementing the lighter printing done in the past. Third, we simplified the surveying question to make it clear to our participants that we are looking for an average impression on the severity of their symptoms rather than the impact: “*So far, what is the severity of your symptoms?*”. We let P1-4 use the first version of the diary for the last four weeks of the first 6-week monitoring period since we did not consider these changes significant.

On the second 6-week cycle, we reintroduced P5 and recruited P6, P7 (Point K in Figure 2). Then, we deployed the revised version of the paper diary to P1-7 (Point L in Figure 2). After 15 days, we interviewed P5 (for a second time), P6 and P7 as we did with P1-4 (Point M in Figure 2).

RESULTS

We report compliance based on Liao et al. [29] guidelines since the diary implements an EMA-like strategy (Table 1). The study length was 49 days with a time variable, self-initiated, once-a-day prompting schedule, and compliance defined as the percentage of days with one answered entry vs study length. Back-filling influenced compliance but it was allowed and a desired feature in our study as it is common practice in traditional Parkinson’s clinical assessment. Although P1-4 used the diary for 12 weeks, we report the first 49 days for consistency as this is the period over which P1-7 participated. For completeness, we also provide the three main symptoms chosen by each participant, the distribution of responses of the seven diaries and the distribution of days with 1,2 or 3 answered entries in Multimedia Appendix A.

Participant	Compliance
P1	91.84%
P2-6	100%
P7	97.95%
Overall	98.54% (SD=3.05; Range=91.84-100)

Table 1. Diary compliance over 49 days

We emphasise that although the symptom survey used in the diary has high face validity, its clinical validity has not been tested. It should be taken as a reference of the level of participant demand we required in the diary since self-reported multiple-answer questions about PD symptoms are normal in clinical assessment. Therefore, although the survey might be suitable for our project’s longitudinal analysis, its outcome values are self-reported and subject to bias. It should not be considered an instrument ready to deploy in clinical settings, but an opportunity to explore a new reporting artefact.

To better understand our participants’ experience with the diary and shed light on what led to this level of compliance, we transcribed the 15-day interviews verbatim using a paid service and performed a thematic analysis following an inductive semantic approach [6]. For consistency, we only included the second interview of P5. Two researchers independently coded the interviews having a moderate agreement (unweighted kappa=0.77) and discussed any discrepancies to

reach a consensus. We identified seven topics: four related to positive attributes of the diary, one identifying an issue with this approach and two miscellaneous features.

Simplicity

Five participants thought the diary was simple and easy to use. The other two had a neutral opinion on it.

"It's just so easy to complete [...] I think the diary is neat, concise, well put together, well made, it's sturdy, not going to fall apart easily, it's well marked out." (P4)

"It's very simple. Very easy to complete. I don't think I've had much difficulty, deciding what ranking to give each measure." (P7)

Considerations for writing

Five participants highlighted at least one of the two considerations we had in mind to support and offset the effect of their tremor when using the diary. One: they did not have to hand-write anything down in the diary. And two: the retractable ink gel pen we provided allowed them to mark the paper using less pressure compared to a ballpoint pen.

"A good feature is that I don't have to write, I can just circle the time and fill in a circle for the symptoms." (P1)

"Oh I like the pen, I like the pen very much. It's just very smooth sort of writing, easy writing, I don't need to press too hard. I think it's better than a ballpen, you don't need as much pressure to write with it which helps with my writing. One of my problems is trouble writing because it's the right hand I've got my tremor in." (P4)

Flexibility

Five participants considered the diary flexible and benefited from it. They talked about flexibility regarding completion time and error detection and correction. Since there was not a fixed time window to answer the diary, they could do it whenever it was more convenient, including times when they would answer it retrospectively the next day after they forgot to do it the night before. Although this might be an issue for other studies and diaries like those tracking medication's ON and OFF cycles, for us, it was acceptable and indeed a desired feature. Also, participants appreciated the fact that each page was marked with the date and therefore allowed them to have a clear view of the day to be reported and at the same time identify or even correct erroneous entries.

"I don't need to worry if I miss it, I can always go back to it, I can just sort of think back if I missed a day thinking, oh what did I do yesterday? [...] sometimes I've missed a day and come back to that the day after." (P4)

"It's easy, one or two days, or one day last week, I was feeling absolutely tired out, and I'd put a three mark in, and I crossed it off and put a four in, because I was really low on energy." (P6)

Implicit reminders

All participants use the physical presence of the diary in a particular place of their homes as an implicit reminder for its

completion. Since we only required them to answer the diary once a day, some people kept their diaries near their bed or where they charge their smartphones. In addition, those participants that wanted to fill in the diary more often kept it in the kitchen where meal or medication time would prompt them to complete it. In fact, reminding through physical presence was described before by Sellen and Harper as an affordance of paper where, for example, employees would be reminded of actions that needed to be taken through the physical presence of documents on their desk or drawers working as a "continuous yet relatively unobtrusive" hint [47].

"I keep it [the diary] on the chest of drawers in the hall, because that's where I plug in and recharge my phone every evening. I don't put it [the diary] in the drawer, I actually have it on the surface and that way I can give myself maximum opportunity to make sure I put something in the diary before I go upstairs to bed." (P3)

"I normally have it with me at my work desk. I think the visual appearance of the diary as I go to bed, where I have it on my desk, is a prompt to me to do something with it when I need to. There's something about having a visual prompt to do it. I'm afraid I've got to that age, where I do that with all sorts of things." (P7)

Interpretation problems

There were two interpretation issues. There was some confusion regarding the scale for the "Low energy" symptom. Participants were asked to rate each symptom from "None" to "High" severity. It was unclear to P2 and P6 whether "High" on the scale for this symptom indicated high energy levels or a high severity of low energy. Both concluded it was the latter. On the other hand, P1, P3, P4 and P5 were not sure if they should record the time when they were filling the diary in or the time at which the reported severity level was present. We emphasised we were looking for an average severity level of each symptom up to the point of logging and therefore instruct them to record the time of reflection.

"It creates a bit of a confusion potentially between whether I'm reporting the section on the day or whether I'm reporting what's happening at the time of recording. So, what I've recorded, I've done it as a reflection of the day, rather than a reflection of that moment." (P3)

The second problem might build a case for the acceptance of in-situ self-reporting of symptoms severity. However, as mentioned in the Background, paper diaries that have been implemented for intensive time-based sampling in Parkinson's are not meant to be used for more than a few days. Technology-based EMA could be an alternative but since interruption burden is also a limitation, it is necessary to overcome two issues: the amount and length of interruption and the friction of accessing the reporting artefact [22]. For example, the use of micro-interactions in smartwatches showed higher compliance compared to a smartphone-based EMA intervention over four weeks [22]. What is more, a custom-made wrist Bluetooth button provided a convenient method for a person with post-traumatic stress disorder to report one symptom leading to hyper-arousal for four months [26]. Likewise, a custom-made

Bluetooth artefact allowed participants to report physical activity or sleep/stress using a touch-sensitive surface for one week [41]. In the context of Parkinson's, these ideas could be adapted for in the wild longitudinal self-reporting. Nevertheless, researchers should explore the role of motor and non-motor symptoms on the acceptability and efficacy of such hardware, in addition to its clinical validity, in case it is meant to be deployed in clinical settings. Whether digital or analogue artefacts are used, the design insights we provide in the Discussion represent actionable items for designers and scientists that can foster creativity and innovation.

Suggesting additional notes

Three participants also suggested having a space for notes describing what was happening when they logged their symptoms. From the beginning of the diary design, we thought that asking something like this would be too demanding for all of our participants, but those who suggested it were keen to do it out of motivation to enrich our analysis and also as a guide to discover patterns in their daily life. These notes are not relevant for our parent study, but researchers interested in collecting them might find it useful to know that our participants considered it doable to write a one word/sentence note per entry in the diary. It is also worth highlighting the case of P7, who developed his own note-taking system within our diary, marking in blue everyday activity labels and in red severe circumstances that affected his day (Figure 8).

"I wonder whether there should be a little spot for an explanation of what the problem is, rather than just filling in the blanks [...] just put on like a one liner about what the problem was." (P4)

Self-awareness of symptoms

Finally, we noticed four participants started being more aware of their symptoms. P1 felt compelled to fill in extra entries in the diary when his symptoms were severe, while P3 confirmed her notion that her symptoms do not fluctuate strongly within days. P6 tried to identify what activities had the worst effect on his symptoms using his daily notes system and hinted at using this information to manage his daily routine.

"It shows some days are better than others. It can pinpoint, associated with the notes that I'm writing, of what is causing me a bad day, and avoid it if possible, or acknowledge that doing this or that, is going to bring a high score on the symptoms." (P6)

Finally, P5 in contrast to P1 was motivated to report her gait when it was better than usual.

"Yesterday I went for a walk and I was quite proud of my walking, I've scored myself because I was thinking, this is good. It is like I'm trying to show how good I can be, to compete with myself." (P5)

She added that this practice is a coping strategy to try to improve her gait and at the same time show others that she is doing well. What is more, since she considers this idea inherent to her, she believes it has been -and will be- consistent.

DISCUSSION

Our paper diary reached almost full compliance over 49 days of use among seven PwP. We provide design implications that support the creation of future artefacts to collect self-reported data in Parkinson's Disease if researchers decide to go back to analogue or to use a digital medium. We reflect on the seven constraints and attributes that guided the diary design and on the seven themes we identified from our participant's experience using such an approach. However, the following implications should not be considered canon for this particular type of diary and must be scrutinised under our study's limitations (see Appendix, Project Constraints): specific hardware requirements, limited funding, a small sample of people with mild Parkinson's and a lack of a formal comparison between the electronic and analogue tools we created.

Reduce participant completion demand. We confirmed that by using a less demanding end-of-day diary we accomplished higher compliance than reported before as suggested by [7, 16]. In our diary, three factors lessened patients' load: having an artefact with a physical footprint that enabled implicit reminders, a flexible completion strategy that adapted to people's routines and allowed them to correct errors, and the "Immediacy of capture" affordance that comes with the use of a pen and allows people to capture their symptoms quickly [46]. Indeed, people with Parkinson's have been shown to be responsive to graspable objects [13].

Design to offset the effect of tremor on input. Our participants appreciated the fact that the diary could be completed without writing and with as little movement of the hand as possible. We made this possible by using questions with ordinal answers, without open-ended inquiries and by pre-printing as much information as possible (i.e. the date). Furthermore, the strokes that were necessary were made with an ink gel pen which our participants perceive as better than a ballpoint pen confirming the anecdotal evidence we had at the start of our study. Interestingly those were also the only people that chose tremor as one of their main symptoms to track with the diary. However, people with Parkinson's with advanced dyskinesias, moderate kinetic tremor (between 3 and 10cm) or rigidity might have problems handling a pen, and in this case, increasing the size of the answer areas or using a marker might allow participants to answer the diary using bolder strokes, or favouring technology might be a good alternative, as it can correct for involuntary movements [28]. For reference, our sample of participants had mild motor symptoms as measured by Part III of the MDS-UPDRS (mean=36 points out of 132, sd=3.53), and none to slight dyskinesias and ON/OFF fluctuations as measured by Part IV of the same scale (avg=3.7 out of 36, sd=5.6). Moreover, based on the conversations we had with our participants every six weeks for the past six months, we know that when their symptoms were particularly severe, they would try to defer some of their daily activities which could include the diary's completion, as they can predict the wearing-off of their medication. Yet, as Parkinson's progresses the percentage of predictable drug cycles decreases and thus researchers might want to test the use of the diary during the worst periods to adapt its design and improve compliance.

Enable implicit reminders All participants used the diary's presence as a trigger to log their symptoms' severity, something we did not expect. Of course, this is not as relevant for diary designs that require a fixed logging time or those that are event-based where the monitored phenomenon is the reminder. Although paper by itself allows for this behaviour, researchers can design other noticeable analogue artefacts that integrate into people's pre-existing habits by being in frequented places. Such devices do not ask people to change their routines but function as an aide memoire. Furthermore, this attribute could be emulated in electronic approaches running on smartphones or computers without recurring to intrusive reminders like alarms, but instead using direct access icons placed on the device's home screen, an ongoing notification on their lock screen, or even with physical objects that participants can place in their environment as with the paper diary.

Design for positive and negative consequences of increased awareness. To the best of our knowledge no research has explored the effect of symptoms awareness on the health of individuals with mild Parkinson's Disease. In the context of self-monitoring, Wilde et al. [54] stated based on previous work that "awareness grows over time if people think what they experience [and] see the relationship to their disease or condition" which was what was observed. The authors established awareness of symptoms as one of two core components of self-monitoring that can lead to improved disease self-management by recognising symptoms and common triggers, and setting and refining goals that can be related to preventing acute illnesses. In fact, P6 identified daily triggers to plan his routine. This is in line with what Ayobi et al. observed when people self-monitored Multiple Sclerosis using different artefacts which helped them to understand (especially to newly diagnosed people) their bodily reactions in everyday life [3]. Nevertheless, as seen in other conditions, there could also be unexpected negative consequences such as participants becoming worried, obsessed or anxious over their symptoms [3, 38] or symptoms (coughing) getting worst after thinking about them [38]. Following this evidence, we also contemplate the possibility where PwP might adhere to their medication intake schedule more faithfully to boost their efficacy but, more worryingly and on an extreme case, this same behaviour could prompt patients to modify it without consulting their doctor. Based on a systematic review of self-care technologies on chronic conditions where the authors recommend fostering reflection by making health and contextual information available [38], we speculate that giving participants access to their diary completion history and having contextual notes about everyday activities increased their awareness and therefore the likelihood of any of the consequences listed above (although we do not know to what degree). Thus, researchers who want to limit this effect could prevent participants from annotating their diaries if it fits their research goals, use shorter diaries and exchange a used one for a new one more often, or use electronic approaches where the control of both circumstances is more granular (e.g., not allowing any retrospective inspection of answers). Besides, other authors should also consider possible bias with reporting strategies like those from P3 and P5, where the severity of symptoms can skew the frequency of diary completion.

Consider the effects of handwritten notes in compliance, encoding burden and data quality. Handwritten notes might provide rich insights from participants self-reporting experience. Likewise, people might find it easier to add notes using pen and paper thanks to the pen affordance of "Integrating information in context" [46]. However, it could also represent an unbearable task to other people; i.e. the cognitive load necessary to generate such notes contributed to the withdrawal of P5 in our study. Furthermore, the availability of contextual notes can be linked to an increase of symptoms awareness as explained above. In any case, researchers should have in mind that transcribing these notes requires either the time to do it manually or the expertise to develop a piece of software that does it automatically. Additionally, prompting for handwritten notes using open questions brings the need for extensive coding and larger item non-response [45].

CONCLUSION

During the agile prototyping of the different self-reporting tools described above, we found out that pen and paper are a suitable method to collect longitudinal day-to-day fluctuations. P1-7's feedback provided us with a detailed view of the problems and virtues of our prototypes that were the base to conduct each redesign.

Although our decision process was tied up to a set of constraints, a button-based or app-based tool might be the right choice under different circumstances [8]. For example, when the chosen device does not interfere with other aspects of the monitoring study. Therefore, in line with other authors, we do not advocate the use of either analogue or electronic approaches as a silver bullet, but rather recommend tailoring either option to the study's goals.

We highlight the attributes that favour our paper diary compared to some technologies. It is cheap at £3.5 per unit, accessible, frictionless, personalised, portable, low-demand, automatically encoded, straightforward and flexible. These are all characteristics that make it a suitable tool for ground truth data collection in our parent study.

APPENDIX

PROJECT CONSTRAINTS

Our design process was delimited by the following constraints. Per our ethics approved protocol, we had three six-week design cycles with an equal number of sessions to deploy a new prototype or get feedback from our users. We had a budget to develop or buy technology of approximately £30 per participant. Participants' smartphones were upgraded to Android 6.1 as a result of a technical fault that forbade continuous Bluetooth sensing in Android 5.1.1 on a Moto X Play device using the Aware Framework. In consequence, any technology-based developed prototype had to be compatible with Android 6.1 and the chipset Qualcomm MSM8939 inside the Moto X Play.

ACKNOWLEDGMENTS

This project is sponsored by the National Council of Science and Technology (CONACyT) and the Secretariat of Public Education (SEP) of Mexico. We also acknowledge the support of the Medical Research Council Confidence in Concept scheme.

REFERENCES

1. Angelo Antonini, Pablo Martinez-Martin, Ray K. Chaudhuri, Marcelo Merello, Robert Hauser, Regina Katzenschlager, Per Odin, Mark Stacy, Fabrizio Stocchi, Werner Poewe, Oliver Rascol, Cristina Sampaio, Anette Schrag, Glenn T. Stebbins, and Christopher G. Goetz. 2011. Wearing-off scales in Parkinson's disease: Critique and recommendations. *Movement Disorders* 26, 12 (10 2011), 2169–2175. DOI: <http://dx.doi.org/10.1002/mds.23875>
2. A. Ashburn, E. Stack, C. Ballinger, L. Fazakarley, and C. Fitton. 2008. The circumstances of falls among people with Parkinson's disease and the use of Falls Diaries to facilitate reporting. *Disability and Rehabilitation* 30, 16 (1 2008), 1205–1212. DOI: <http://dx.doi.org/10.1080/09638280701828930>
3. Amid Ayobi, Paul Marshall, Anna L. Cox, and Yunan Chen. 2017. Quantifying the Body and Caring for the Mind. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, New York, New York, USA, 6889–6901. DOI: <http://dx.doi.org/10.1145/3025453.3025869>
4. Niall Bolger, Angelina Davis, and Eshkol Rafaeli. 2003. Diary Methods: Capturing Life as it is Lived. *Annual Review of Psychology* 54, 1 (2 2003), 579–616. DOI: <http://dx.doi.org/10.1146/annurev.psych.54.101601.145030>
5. Niall Bolger, Patrick E Shrout, Amie S Green, Eshkol Rafaeli, and Harry T Reis. 2006. Paper or plastic revisited: Let's keep them both—Reply to Broderick and Stone (2006); Tennen, Affleck, Coyne, Larsen, and DeLongis (2006); and Takarangi, Garry, and Loftus (2006). (2006).
6. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
7. Joan E. Broderick and Arthur A. Stone. 2006. Paper and electronic diaries: Too early for conclusions on compliance rates and their effects—Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods* 11, 1 (3 2006), 106–111. DOI: <http://dx.doi.org/10.1037/1082-989X.11.1.106>
8. Michelle Clare Carter, Victoria Jane Burley, Camilla Nykjaer, and Janet Elizabeth Cade. 2013. Adherence to a Smartphone Application for Weight Loss Compared to Website and Paper Diary: Pilot Randomized Controlled Trial. *Journal of Medical Internet Research* 15, 4 (4 2013), e32. DOI: <http://dx.doi.org/10.2196/jmir.2283>
9. Tamlin Conner Christensen, Lisa Feldman Barrett, Eliza Bliss-Moreau, Kirsten Lebo, and Tamlin Conner Christensen. 2003. A Practical Guide to Experience-Sampling Procedures. *Journal of Happiness Studies* 4, 1 (2003), 53–78. DOI: <http://dx.doi.org/10.1023/A:1023609306024>
10. Oystein Dale and Kaare Birger Hagen. 2007. Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data. *Journal of Clinical Epidemiology* 60, 1 (1 2007), 8–17. DOI: <http://dx.doi.org/10.1016/j.jclinepi.2006.04.005>
11. Denzil Ferreira, Vassilis Kostakos, and Anind K. Dey. 2015. AWARE: Mobile Context Instrumentation Framework. *Frontiers in ICT* 2 (4 2015). DOI: <http://dx.doi.org/10.3389/fict.2015.00006>
12. Linda A. Fletcher, Darin J. Erickson, Traci L. Toomey, and Alexander C. Wagenaar. 2003. Handheld Computers. *Evaluation Review* 27, 2 (4 2003), 165–178. DOI: <http://dx.doi.org/10.1177/0193841X02250527>
13. Adam Galpin, Steven P. Tipper, Jeremy P.R. Dick, and Ellen Poliakoff. 2011. Object affordance and spatial-compatibility effects in Parkinson's disease. *Cortex* 47, 3 (2011), 332 – 341. DOI: <http://dx.doi.org/https://doi.org/10.1016/j.cortex.2010.01.011>
14. Alan Godfrey, Alan Bourke, Silvia Del Din, Rosie Morris, Aodhan Hickey, Jorunn L. Helbostad, and Lynn Rochester. 2016. Towards holistic free-living assessment in Parkinson's disease: Unification of gait and fall algorithms with a single accelerometer. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 651–654. DOI: <http://dx.doi.org/10.1109/EMBC.2016.7590786>
15. Christopher G Goetz, Barbara C Tilley, Stephanie R Shaftman, Glenn T Stebbins, Stanley Fahn, Pablo Martinez-Martin, Werner Poewe, Cristina Sampaio, Matthew B Stern, Richard Dodel, Bruno Dubois, Robert Holloway, Joseph Jankovic, Jaime Kulisevsky, Anthony E Lang, Andrew Lees, Sue Leurgans, Peter a LeWitt, David Nyenhuis, C Warren Olanow, Olivier Rascol, Anette Schrag, Jeanne a Teresi, Jacobus J van Hilten, and Nancy LaPelle. 2008. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Movement disorders : official journal of the Movement Disorder Society* 23, 15 (11 2008), 2129–70. DOI: <http://dx.doi.org/10.1002/mds.22340>
16. Amie S. Green, Eshkol Rafaeli, Niall Bolger, Patrick E. Shrout, and Harry T. Reis. 2006. Paper or plastic? Data equivalence in paper and electronic diaries. *Psychological Methods* 11, 1 (2006), 87–105. DOI: <http://dx.doi.org/10.1037/1082-989X.11.1.87>
17. Robert A. Hauser, Frieda Deckers, and Philippe Leheret. 2004. Parkinson's disease home diary: Further validation and implications for clinical trials. *Movement Disorders* 19, 12 (12 2004), 1409–1413. DOI: <http://dx.doi.org/10.1002/mds.20248>
18. Turid Heiberg, Tore K Kvien, Øystein Dale, Petter Mowinckel, Gerd J Aanerud, Ann B Songe-Møller, Till Uhlig, and Kåre B Hagen. 2007. Daily health status registration (patient diary) in patients with rheumatoid arthritis: A comparison between personal digital assistant and paper-pencil format. *Arthritis Care & Research* 57, 3 (2007), 454–460.

19. Sharpley Hsieh, Samantha Schubert, Christopher Hoon, Eneida Mioshi, and John R Hodges. 2013. Validation of the Addenbrooke's Cognitive Examination III in frontotemporal dementia and Alzheimer's disease. *Dementia and geriatric cognitive disorders* 36, 3-4 (2013), 242–250.
20. Heather Hunter, Lynn Rochester, Rosie Morris, and Sue Lord. 2017. Longitudinal falls data in Parkinson's disease: feasibility of fall diaries and effect of attrition. *Disability and Rehabilitation* (6 2017), 1–6. DOI : <http://dx.doi.org/10.1080/09638288.2017.1329357>
21. M E Hyland, C A Kenyon, R Allen, and P Howarth. 1993. Diary keeping in asthma: comparison of written and electronic methods. *BMJ (Clinical research ed.)* 306, 6876 (2 1993), 487–9. <http://www.ncbi.nlm.nih.gov/pubmed/8448458>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1676776>
22. Stephen Intille, Caitlin Haynes, Dharam Maniar, Aditya Ponnada, and Justin Manjourides. 2016. μ EMA. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16*. ACM Press, New York, New York, USA, 1124–1128. DOI : <http://dx.doi.org/10.1145/2971648.2971717>
23. T K Kvien, P Mowinckel, T Heiberg, K L Dammann, Ø Dale, G J Aanerud, T N Alme, and T Uhlig. 2005. Performance of health status measures with a pen based personal digital assistant. *Annals of the Rheumatic Diseases* 64, 10 (10 2005), 1480–1484. DOI : <http://dx.doi.org/10.1136/ard.2004.030437>
24. Jessica Lam, Ronald G. Barr, Nicole Catherine, Haley Tsui, Claire L. Hahnhaussen, Julie Pauwels, and Rollin Brant. 2010. Electronic and Paper Diary Recording of Infant and Caregiver Behaviors. *Journal of Developmental & Behavioral Pediatrics* 31, 9 (10 2010), 1. DOI : <http://dx.doi.org/10.1097/DBP.0b013e3181e416ae>
25. Shannon J Lane, Nancy M Heddle, Emmy Arnold, and Irwin Walker. 2006. A review of randomized controlled trials comparing the effectiveness of hand held computers with paper methods for data collection. *BMC medical informatics and decision making* 6 (5 2006), 23. DOI : <http://dx.doi.org/10.1186/1472-6947-6-23>
26. Jakob Eg Larsen, Kasper Eskelund, and Thomas Blomseth Christiansen. 2017. Active Self-Tracking of Subjective Experience with a One-Button Wearable: A Case Study in Military PTSD. *arXiv preprint arXiv:1703.03437* (2017).
27. Jean-Philippe Laurenceau and Niall Bolger. 2005. Using Diary Methods to Study Marital and Family Processes. *Journal of Family Psychology* 19, 1 (3 2005), 86–97. DOI : <http://dx.doi.org/10.1037/0893-3200.19.1.86>
28. James L Levine and Michael A Schappert. 2005. A mouse adapter for people with hand tremor. *IBM Systems Journal* 44, 3 (2005), 621–628.
29. Yue Liao, Kara Skelton, Genevieve Dunton, and Meg Bruening. 2016. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE checklist for reporting EMA studies (CREMAS). *Journal of medical Internet research* 18, 6 (2016).
30. Paul Luff, Christian Heath, Moira Norrie, Beat Signer, and Peter Herdman. 2004. Only touching the surface. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work - CSCW '04*. ACM Press, New York, New York, USA, 523. DOI : <http://dx.doi.org/10.1145/1031607.1031695>
31. Heikki Luukinen, Keijo Koski, Liisa Hiltunen, and Sirkka-Liisa Kivelä. 1994. Incidence rate of falls in an aged population in Northern Finland. *Journal of Clinical Epidemiology* 47, 8 (8 1994), 843–850. DOI : [http://dx.doi.org/10.1016/0895-4356\(94\)90187-2](http://dx.doi.org/10.1016/0895-4356(94)90187-2)
32. Kelly E. Lyons and Rajesh Pahwa. 2007. Electronic motor function diary for patients with Parkinson's disease: A feasibility study. *Parkinsonism & Related Disorders* 13, 5 (7 2007), 304–307. DOI : <http://dx.doi.org/10.1016/j.parkreldis.2006.06.008>
33. Altaf Mangera, Alberto Marzo, Nicola Heron, Dayan Fernando, Khawar Hameed, Abdel-Hamid A. Soliman, Mike Bradley, Ian Hosking, Mohamed Abdel-Maguid, Martin Levermore, Wendy B. Tindale, and Christopher Chapple. 2014. Development of two electronic bladder diaries: A patient and healthcare professionals pilot study. *Neurourology and Urodynamics* 33, 7 (9 2014), 1101–1109. DOI : <http://dx.doi.org/10.1002/nau.22469>
34. Johan Marinus, Martine Visser, Anne M Stiggelbout, Jose Martin Rabey, Ubaldo Bonuccelli, Peter H Kraus, and Jacobus Bob J van Hilten. 2002. Activity-based diary for Parkinson's disease. *Clinical neuropharmacology* 25, 1 (2002), 43–50.
35. Rob McCarney, James Warner, Steve Iliffe, Robbert van Haselen, Mark Griffin, and Peter Fisher. 2007. The Hawthorne Effect: a randomised, controlled trial. *BMC Medical Research Methodology* 7, 1 (12 2007), 30. DOI : <http://dx.doi.org/10.1186/1471-2288-7-30>
36. Mattijn Morren, Sandra Dulmen, Jessika Ouwerkerk, and Jozien Bensing. 2009. Compliance with momentary pain measurement using electronic diaries: A systematic review. *European Journal of Pain* 13, 4 (4 2009), 354–365. DOI : <http://dx.doi.org/10.1016/j.ejpain.2008.05.010>
37. Jan M. Noyes and Kate J. Garland. 2008. Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics* 51, 9 (9 2008), 1352–1375. DOI : <http://dx.doi.org/10.1080/00140130802170387>
38. Francisco Nunes, Nervo Verdezoto, Geraldine Fitzpatrick, Morten Kyng, Erik Grönvall, and Cristiano Storni. 2015. Self-Care Technologies in HCI. *ACM Transactions on Computer-Human Interaction* 22, 6 (12 2015), 1–45. DOI : <http://dx.doi.org/10.1145/2803173>

39. Dag Nyholm, Jan Kowalski, and Sten-Magnus Aquilonius. 2004. Wireless real-time electronic data capture for self-assessment of motor function and quality of life in Parkinson's disease. *Movement Disorders* 19, 4 (4 2004), 446–451. DOI: <http://dx.doi.org/10.1002/mds.10690>
40. Spyridon Spyros Papapetropoulos. 2012. Patient Diaries As a Clinical Endpoint in Parkinson's Disease Clinical Trials. *CNS Neuroscience & Therapeutics* 18, 5 (5 2012), 380–387. DOI: <http://dx.doi.org/10.1111/j.1755-5949.2011.00253.x>
41. Gaurav Paruthi, Shriti Raj, Ankita Gupta, Chuan-Che Huang, Yung-Ju Chang, and Mark W Newman. 2017. HEED. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers on - UbiComp '17*. ACM Press, New York, New York, USA, 181–184. DOI: <http://dx.doi.org/10.1145/3123024.3123162>
42. V. Peto, C. Jenkinson, and R. Fitzpatrick. 1998. PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. *Journal of Neurology* 245, S1 (4 1998), S10–S14. DOI: <http://dx.doi.org/10.1007/PL00007730>
43. S L Pullman. 1998. Spiral analysis: a new technique for measuring tremor with a digitizing tablet. *Movement disorders : official journal of the Movement Disorder Society* 13 Suppl 3 (1998), 85–9. <http://www.ncbi.nlm.nih.gov/pubmed/9827601>
44. J Reimer, M Grabowski, O Lindvall, and P Hagell. 2004. Use and interpretation of on/off diaries in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 75, 3 (2004), 396–400. DOI: <http://dx.doi.org/10.1136/jnnp.2003.022780>
45. Urša Reja, Katja Lozar Manfreda, Valentina Hlebec, and Vasja Vehovar. 2003. Open-ended vs. close-ended questions in web questionnaires. *Developments in applied statistics* 19, 1 (2003), 159–177.
46. Yann Riche, Nathalie Henry Riche, Ken Hinckley, Sheri Panabaker, Sarah Fuelling, and Sarah Williams. 2017. As We May Ink?. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, New York, New York, USA, 3241–3253. DOI: <http://dx.doi.org/10.1145/3025453.3025716>
47. Abigail J. Sellen and Richard Harper. 2002. *The myth of the paperless office*. MIT Press. 231 pages.
48. Saul Shiffman, Arthur A Stone, and Michael R Hufford. 2008. Ecological Momentary Assessment. *Annu. Rev. Clin. Psychol* 4 (2008), 1–32. DOI: <http://dx.doi.org/10.1146/annurev.clinpsy.3.022806.091415>
49. Jennifer N Stinson and Jennifer N. 2009. Improving the Assessment of Pediatric Chronic Pain: Harnessing the Potential of Electronic Diaries. *Pain Research and Management* 14, 1 (2009), 59–64. DOI: <http://dx.doi.org/10.1155/2009/915302>
50. Howard Tennen, Glenn Affleck, James C. Coyne, Randy J. Larsen, and Anita DeLongis. 2006. Paper and plastic in daily diary research: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods* 11, 1 (2006), 112–118. DOI: <http://dx.doi.org/10.1037/1082-989X.11.1.112>
51. Kwon Kuen Tsang. 2012. The use of midpoint on Likert Scale: The implications for educational research. *Hong Kong Teachers' Centre Journal* 11, 1 (2012), 121–130.
52. Julio Vega. 2017. JulioV/paperstream: PaperStream. (Nov. 2017). DOI: <http://dx.doi.org/10.5281/zenodo.1048283>
53. G Velikova, E P Wright, A B Smith, A Cull, A Gould, D Forman, T Perren, M Stead, J Brown, and P J Selby. 1999. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 17, 3 (3 1999), 998–1007. DOI: <http://dx.doi.org/10.1200/JCO.1999.17.3.998>
54. Mary H Wilde and Suzanne Garvin. 2007. A concept analysis of self-monitoring. *Journal of Advanced Nursing* 57, 3 (2007), 339–350.
55. Jerome A Yesavage and Javid I Sheikh. 1986. 9/Geriatric depression scale (GDS) recent evidence and development of a shorter version. *Clinical gerontologist* 5, 1-2 (1986), 165–173.